Review by the
Office of Program Evaluation and Government Accountability (OPEGA)
Response from the Office of Information Technology (OIT)
March 1, 2013

TOPIC 3: SUPPORTING THE DATA NEEDS OF EXECUTIVE BRANCH DEPARTMENTS

1. Executive Summary

Executive Outline
> Goals / Strategy (2013 - 2014)
> Roles for Data Reporting
> Timeline (summary)
> Industry Best Practices
> Current State
> Desired State
> Timeline (details)

Goals / Strategy (2013 - 2014)
- Current: 16 departments – all doing differently (no strategy)
- Future: Need leadership / strategy owner:
  - Business intelligence (BI) trends
  - Data mining trends

Roles for Data and Reporting:

Systems Analyst	⟷	**Business Analyst**
OIT Owns		**Agency Owns**
Hosting		Data ownership
Professional staff for support/ consulting		Analysis
Data structure		Interpretation
Tools		Data mining

Timeline (summary):

Mid-2013	End of 2013	Mid-2014	Ongoing
• Survey agencies' understanding of data and analytics requirements. • Through IT Executive Committee, clarify agency and OIT roles for data and reporting.	• Train agency business analysts and OIT systems analysts in data structures and analytical tools.	• Recruit full-time OIT Data Evangelist, working with all agency data stakeholders to continuously increase quality of data and analytics.	• Ensure quality-of-service for OIT Systems Analysts.

2. INDUSTRY BEST PRACTICES

Roles and Skills for Data Analysis and Reporting:
- Business users are the owners of data. IT is the caretaker of that data.
- Business users own analysis of data. IT owns programming of data, as well as selecting and maintaining analytics tools.
- IT trains business users in structure and organization of data, as well how to use analytics tools. Modern analytics tools empower business users with intuitive point-and-click interfaces. Report queries can be constructed and refined by business users, and multiple sources of back-end data can be dragged onto the same screen.
- For every element of business data, there must be an official source-of-truth.

Business users	IT
Owners of data	Caretaker of data
Analysis of data	Programming and selecting analytics tools
Use analytics tools to understand data and develop reports	Train business users in data structure and organization, and how to use the tools

3. CURRENT STATE

- Inconsistent levels of staff and skills for data analysis. Some agencies have staff capable of analyzing their data, but others do not.
- No consistent expectation between OIT and agencies about data ownership versus caretaking roles.
- Inconsistent levels of agencies' understanding of their data and analytics requirements.
- From the OIT standpoint, reporting is just part of application development. If an agency desires to pay for OIT to create its business reports, OIT does so without questioning whether that is the most optimal arrangement.
- Incomplete awareness and documentation of the source-of-truth for business data elements.
- No single owner within OIT evangelizing the increasing importance of data and analytics.

4. DESIRED STATE

- Agencies and OIT both understand their data and analytics requirements.
- Clarify agency and OIT roles for data and reporting.
    - The agency is responsible for analyzing their data.
    - OIT is responsible for maintaining and programming the data, as well as selecting and maintaining analytics tools.
- Every agency / bureau/business unit, depending on size and business need, will assign Business Analysts. They do require a solid understanding of their data from a business perspective.
- OIT will train the Business Analyst in:
    - Analytics tools (OBIEE, IBM COGNOS, SAP Crystal Reports, etc.)
    - Data Structures (i.e., HOW the data are organized). But the actual interrogation and interpretation belongs to the Business Analysts.
- A single OIT Data Evangelist working with all data stakeholders, continuously increasing the quality of data and analytics.
- For each data element, ensure a documented source-of-truth.

5. TIMELINE

- **Mid-2013:**
  - OIT Technology Business Consultants will survey agencies' understanding of their data and analytics requirements. Status update every quarter.
  - Mid-2013: Through IT Executive Committee, clarify agency and OIT roles for data and reporting.

- **End of 2013:**
  - Within 3 months of the agency assigning a Business Analyst, OIT will:
    - Pair the agency Business Analyst with an OIT Systems Analyst.
    - Provide the Business Analyst with analytics tools.
    - Train the Business Analyst in relevant data structures and chosen analytics tools.

- **Mid-2014:**
  - Recruit a full-time OIT Data Evangelist, working with all data stakeholders to continuously increase quality of data and analytics.

- **Ongoing:**
  - Associate CIO for Applications will ensure quality-of-service for OIT Systems Analysts.

Appendix

Business Intelligence (or Analytics, or Data Mining) is a collection of processes and technologies that extract meaningful information out of raw data. This provides historical, current, and predictive patterns and insights, which would otherwise remain hidden. What is commonly called Reporting is part of Business Intelligence.

Traditional Business Intelligence operates on data stored in databases. This is the kind of data that are called "structured" because the raw data have been organized into a pre-designed structure inside the database. While structured data remain immensely useful for exploring patterns and insights, it suffers from the following limitations:
1. The overhead of structuring within the database
2. Ideally suited to data that are primarily, though not exclusively, numerical. Not efficient for free-format texts, images, sounds, etc.
3. Ideally suited to data that do not change too frequently
4. Not efficient for very large volumes of data

Traditional Business Intelligence tools include Oracle Business Intelligence, IBM Cognos, Informatica PowerCenter, etc.

In order to overcome the limitations of structured data analytics, there is now a new area of Business Intelligence called Big Data. The processes and technologies of Big Data are better suited to negotiate data that are unstructured, fast-changing, and in volumes exceeding the capacity of traditional analytics. The key features of Big Data are:
1. Decomposition of the data payload into chunks
2. Parallel, simultaneous analysis of these chunks
3. Integration of the parallel results post-analysis

Big Data Business Intelligence tools include Apache Hadoop, Google MapReduce, etc.